# Hybrid Data Assimilation without Ensemble Filtering

Ricardo Todling and Amal El Akkraoui

Global Modeling and Assimilation Office
NASA

National Centers for Environmental Prediction
16 January 2013

*Contributions from: D. Kleist, D. Parrish, R. Treadon, and J. Whitaker*

# Outline



# Variational Formulations

*FGAT 3dVar-ensemble Hybrid*:

$$J(\delta\mathbf{x}) = \frac{1}{2}\,\delta\mathbf{x}^T\mathbf{B}_h{}^{-1}\delta\mathbf{x} + \frac{1}{2}\sum_{k=1}^{K}[\mathbf{H}_k\delta\mathbf{x} - \mathbf{d}_k]^T\mathbf{R}_k^{-1}[\mathbf{H}_k\delta\mathbf{x} - \mathbf{d}_k] + J_x$$

where

- $\mathbf{B}_h = \beta\mathbf{B} + (1-\beta)\mathbf{B}_e \circ \mathbf{C}$ is a *hybrid* of static and ensemble-based error covariances, $\mathbf{B}$ and $\mathbf{B}_e$ respectively;
- $\mathbf{C}$ is a localization error covariance of compact support;
- the control variable changes to be $\delta\mathbf{x} = \delta\mathbf{x}_0 + \sum_m^M \delta\mathbf{x}_m^e \circ \alpha_m$, for an ensemble with a total of $M$ members $\delta\mathbf{x}_m^e$;
- NCEP and GMAO get $\delta\mathbf{x}_m^e$ by using the EnKF plus inflation.
- NOTE: in 3d-Var, the alpha-control variable augmentation leads to a problem similar to that posed to bias-correct the background fields, the only difference being that the hybrid approach dynamically updates the bias error covariance.

# Variational Formulations

*FGAT 3dVar-ensemble Hybrid*:

$$J(\delta\mathbf{x}) = \frac{1}{2}\,\delta\mathbf{x}^T\mathbf{B}_h{}^{-1}\delta\mathbf{x} + \frac{1}{2}\sum_{k=1}^{K}[\mathbf{H}_k\delta\mathbf{x} - \mathbf{d}_k]^T\mathbf{R}_k^{-1}[\mathbf{H}_k\delta\mathbf{x} - \mathbf{d}_k] + J_x$$

where

- $\mathbf{B}_h = \beta\mathbf{B} + (1-\beta)\mathbf{B}_e \circ \mathbf{C}$ is a *hybrid* of static and ensemble-based error covariances, $\mathbf{B}$ and $\mathbf{B}_e$ respectively;
- $\mathbf{C}$ is a localization error covariance of compact support;
- the control variable changes to be $\delta\mathbf{x} = \delta\mathbf{x}_0 + \sum_m^M \delta\mathbf{x}_m^e \circ \alpha_m$, for an ensemble with a total of $M$ members $\delta\mathbf{x}_m^e$;
- NCEP and GMAO get $\delta\mathbf{x}_m^e$ by using the EnKF plus inflation.
- NOTE: in 3d-Var, the alpha-control variable augmentation leads to a problem similar to that posed to bias-correct the background fields, the only difference being that the hybrid approach dynamically updates the bias error covariance.

# Problem Statement

- Hybrid DA schemes include both multiplicative and additive inflation
- Evaluations in GEOS DAS suggest:
  - Hybrid approach provides noticeable improvements only when using additive inflation, i.e., EnKF alone doesn't do it
  - Forecasts from EnKF analyses plus additive inflation result in mild spread within the background time window
  - It seems that much of the initial (analysis) spread can be simulated with additive inflation alone
  - Appreciable background spread is obtained in the latter case

*Question: how does hybrid-DA perform when the ensemble filter is dropped and an ensemble of analyses is created from simply additively inflating the central analysis?*

# Problem Statement

- Hybrid DA schemes include both multiplicative and additive inflation
- Evaluations in GEOS DAS suggest:
  - Hybrid approach provides noticeable improvements only when using additive inflation, i.e., EnKF alone doesn't do it
  - Forecasts from EnKF analyses plus additive inflation result in mild spread within the background time window
  - It seems that much of the initial (analysis) spread can be simulated with additive inflation alone
  - Appreciable background spread is obtained in the latter case

*Question: how does hybrid-DA perform when the ensemble filter is dropped and an ensemble of analyses is created from simply additively inflating the central analysis?*

# Reasoning behind the filter-free approach

Take the true state to evolve according to: $\mathbf{x}^t = \tilde{\mathbf{m}}(\mathbf{x}_0^t)$
Take the forecasting model to evolve as in: $\mathbf{x}^f = \mathbf{m}(\mathbf{x}_0)$
To first order, an initial uncertainty $\delta\mathbf{e}$ produces to following forecast error:

$$\delta\mathbf{e}^f \approx \mathbf{M}(\delta\mathbf{x}_0 + \delta\mathbf{e}) + \mathbf{q}$$

where $\delta\mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_0^t$ and $q \equiv \mathbf{m}(\mathbf{x}_0^t) - \tilde{\mathbf{m}}(\mathbf{x}_0^t)$.

- For unbiased models, initial uncertainty does not represent model error
- First term in the expression represents propagation of initial uncertainty
- In present ensemble (hybrid) implementations:
  $\delta\mathbf{x}_0$ is a member analysis increment
  $\delta\mathbf{e}$ is an inflation term added to the increment
  $\rightarrow$ These represent redundant treatment of initial uncertainty

*The present work evaluates the case when $\delta\mathbf{x}_0$ is ignored; that is, the ensemble is generated from inflated initial errors over the central analysis.*

## Atmospheric GCM

- Fully ESMF-based
- Cubed-sphere hydrostatic dynamical core
- RAS-Bacmeister convective physics
- Chou-Suarez radiation scheme
- Koster et al. catchment land-surface model
- Lock et al. turbulence physics
- Interactive ozone
- Interactive GOCART aerosols
- OSTIA-prescribed SST

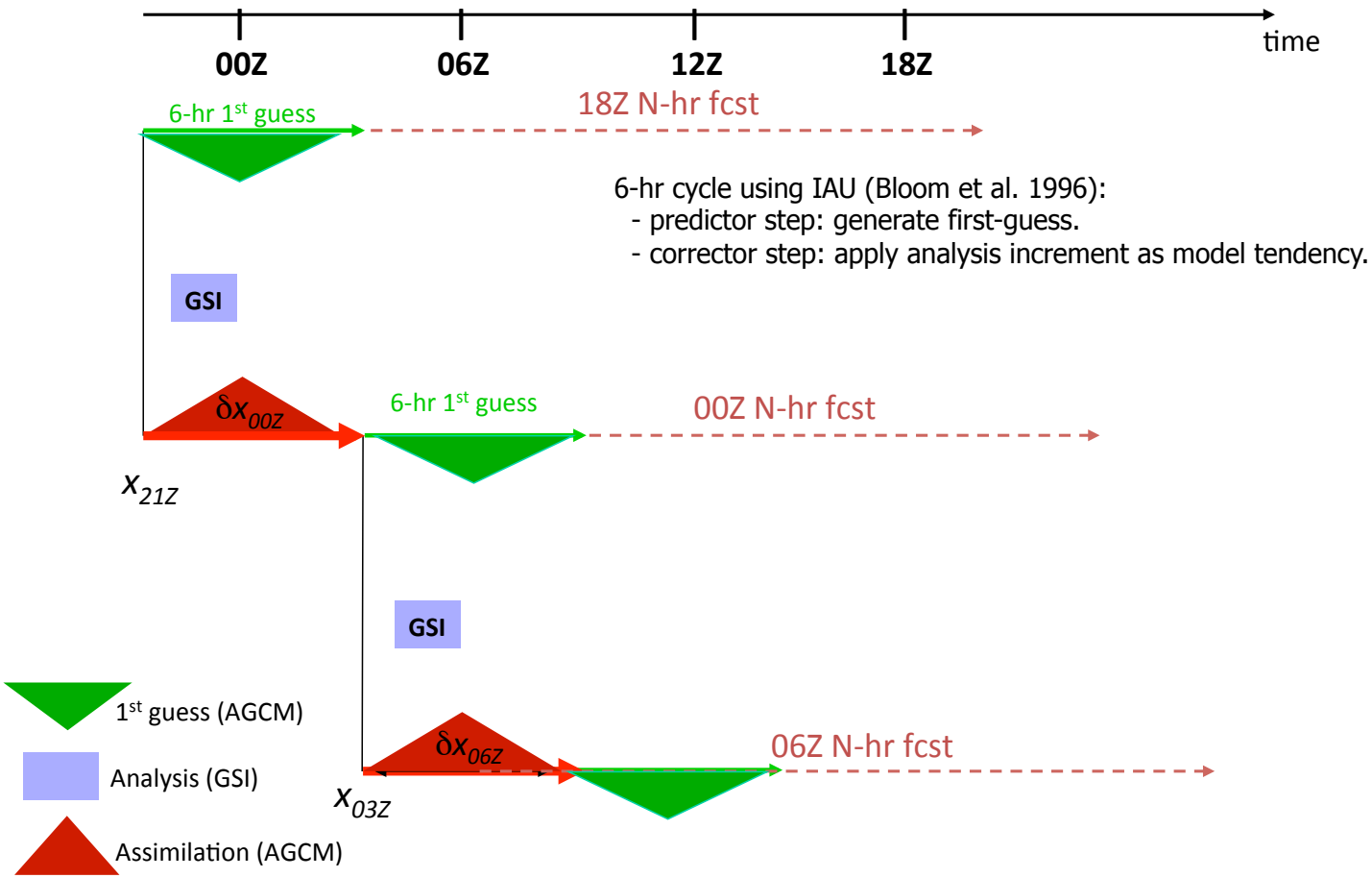
## Analysis: GSI

- FGAT 3D-Var
- IAU-based assimilation
- TLNMC balance
- JCSDA CRTM
- Double-PCG minimization

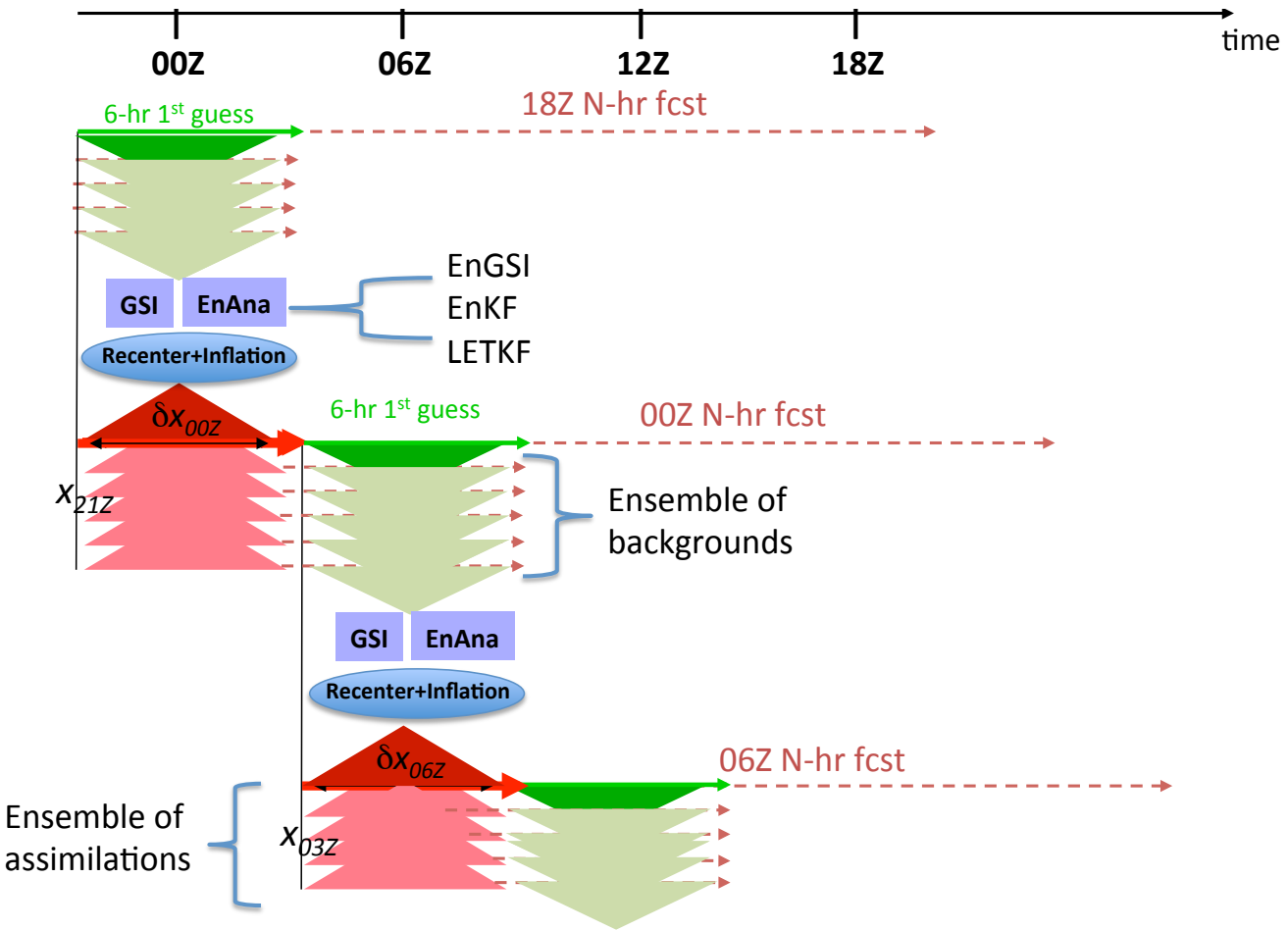
## Ensemble filter

- ESRL-NCEP EnKF
- Full obs but ozone and precip

# Schematic of GEOS IAU-based 3dVar



# Schematic of IAU-based Hybrid 3dVar



## Hybrid Experimental Setting

- Central DAS: $0.5^o$ outer and inner loops; 72-levels
- 32 Ensemble Forecasts: $1.0^o$; 72-levels
- GSI Hybrid/Static **B**: 50% / 50%
- TLNMC applied to both static & hybrid covariances
- Vertical & horizontal localizations applied to ensemble **B**
- Add/ve perturbations scaled from NMC-like 48-24hr forecasts
- Experiment period (after spin up): April 2012

### EnKF

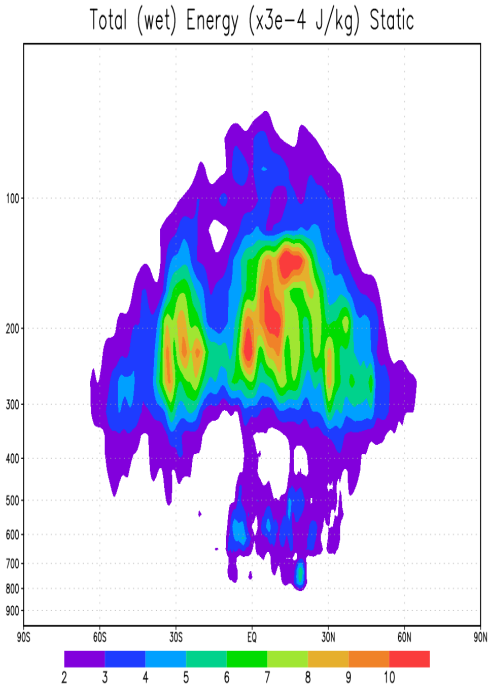
- Additive perturbation: 0.25
- EnKF

### Filter-Free

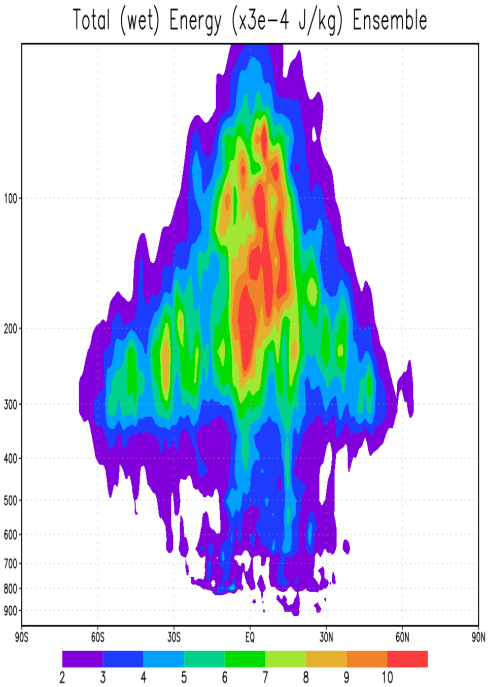
- Additive perturbation: 0.6
- No Ensemble Filtering

# Analysis Increment as Total Energy for 00 UTC on 1 Jun 2012

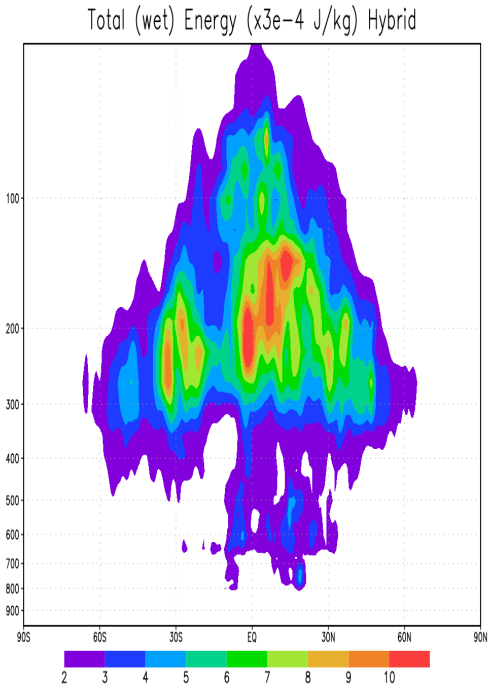
## Static-Only




## 32-mem Ens-Only

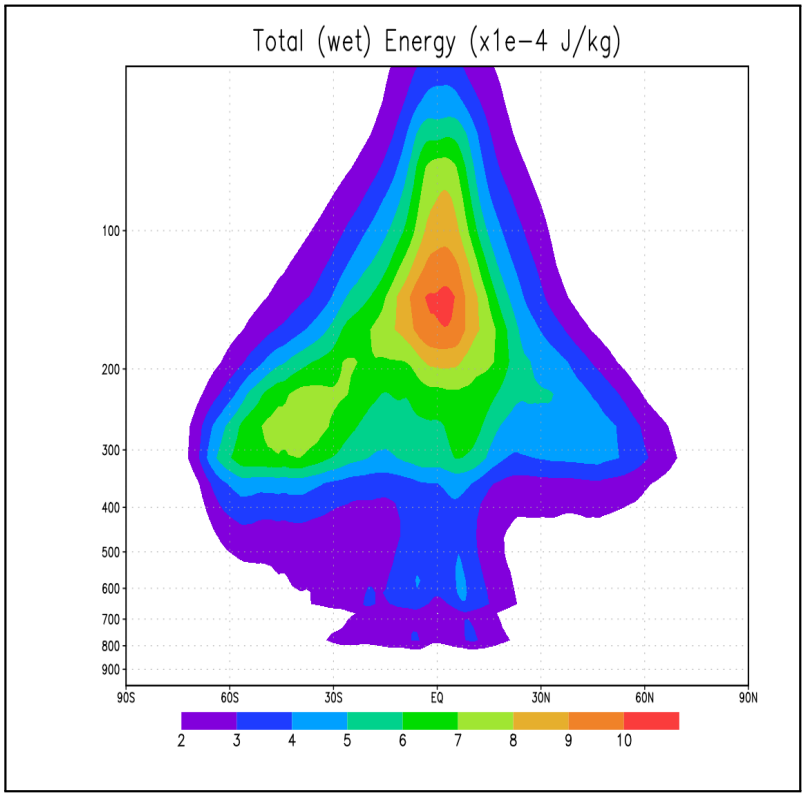



## Hybrid (50%/50%)

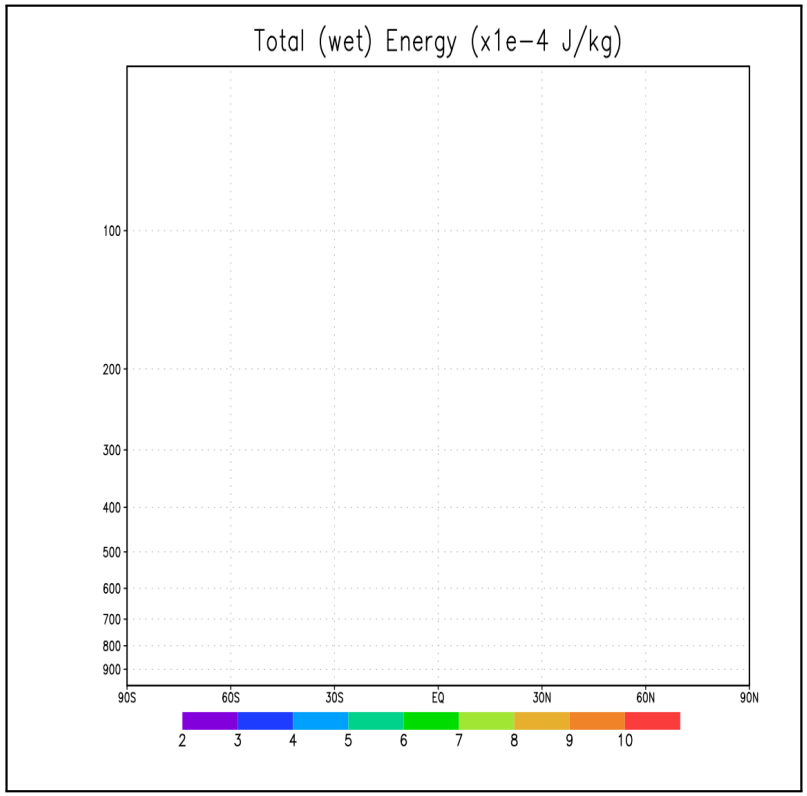


# 0-hr Analyses Spread (before additive inflation) as Total Energy

## EnKF-based hybrid

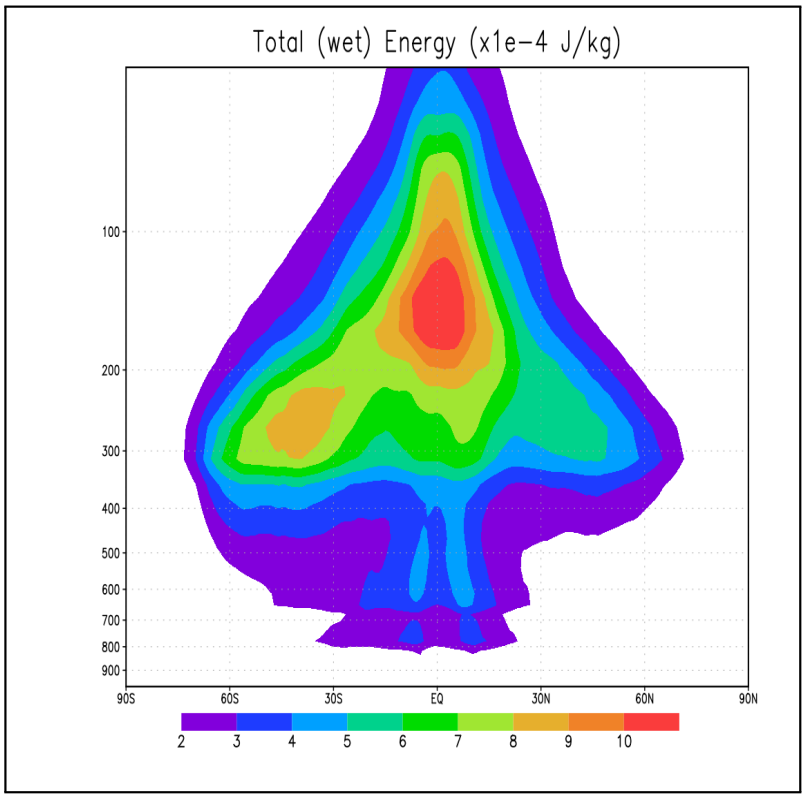



## Filter-Free hybrid

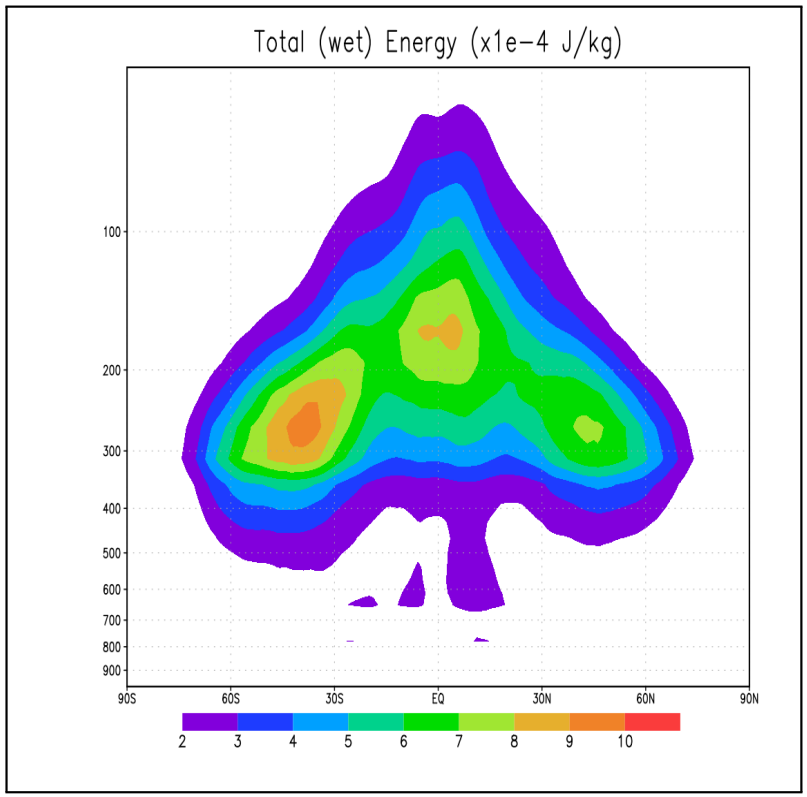


# 3-hr Background Spread

## EnKF-based hybrid

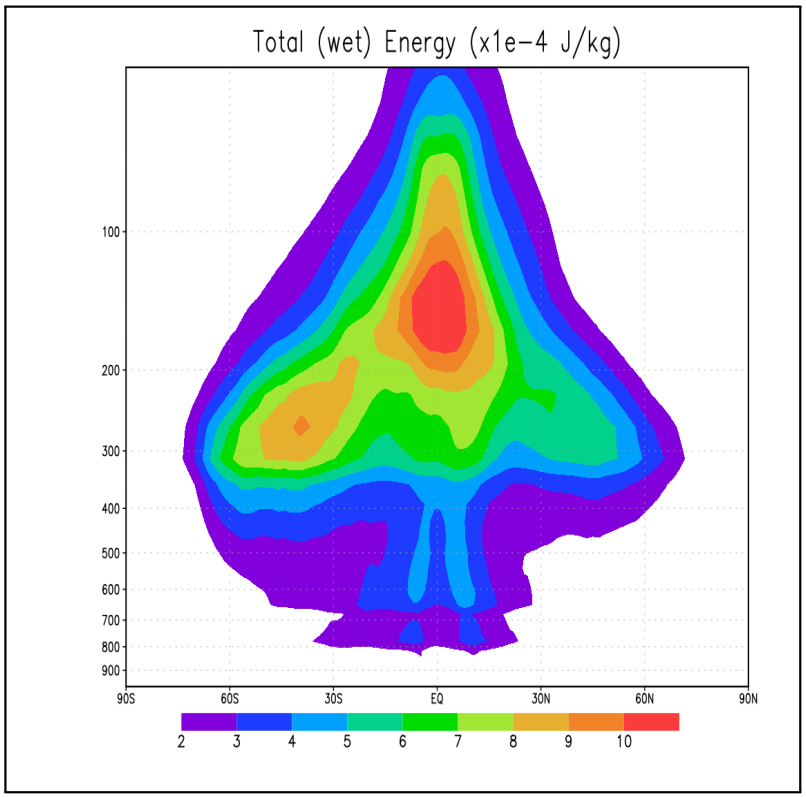



## Filter-Free hybrid

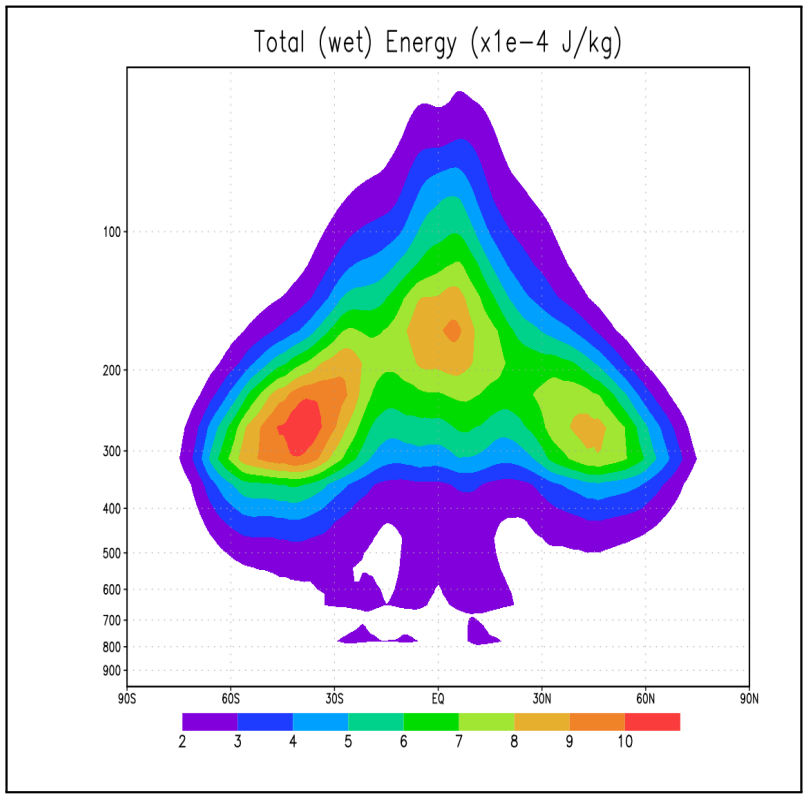


# 6-hr Background Spread

## EnKF-based hybrid

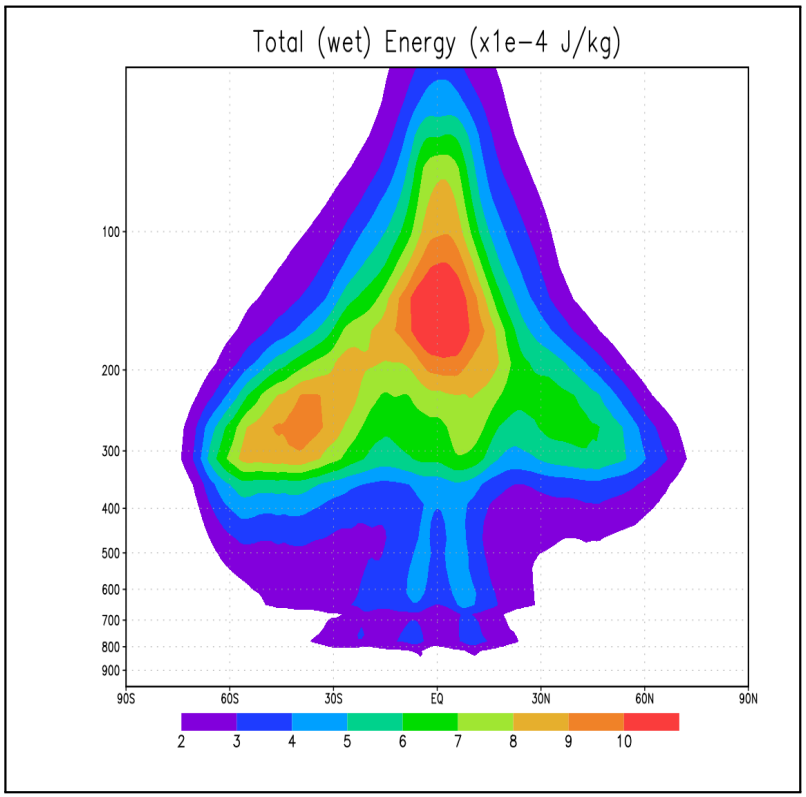



## Filter-Free hybrid

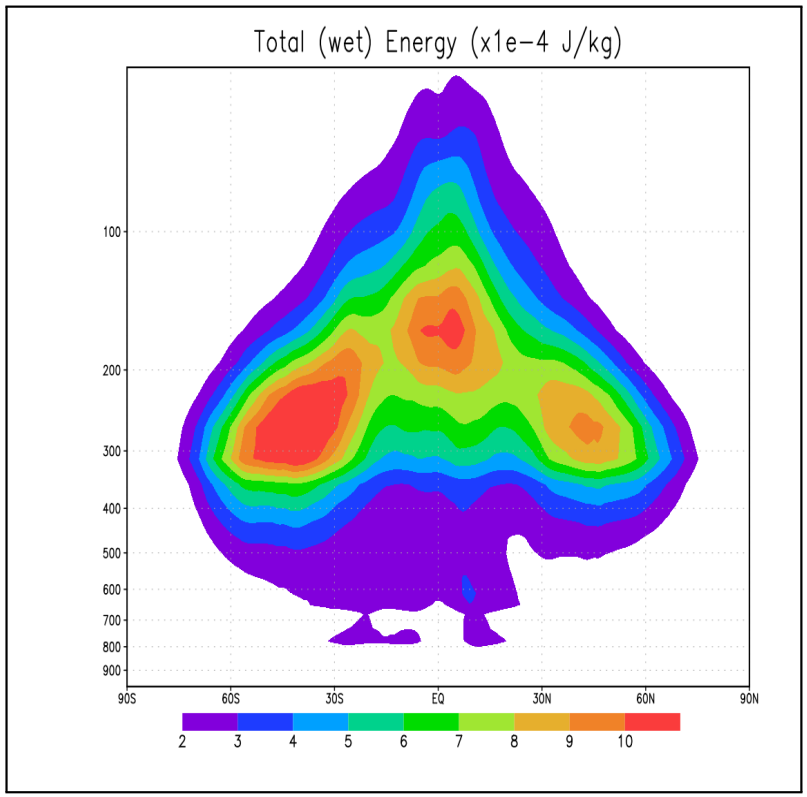


# 9-hr Background Spread

## EnKF-based hybrid

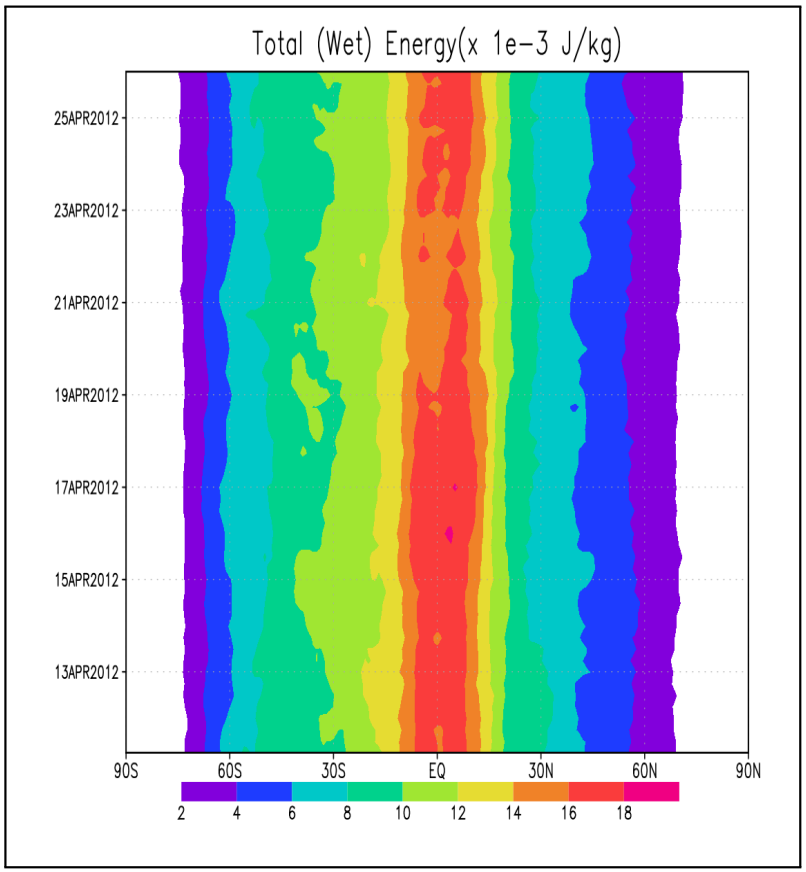



## Filter-Free hybrid

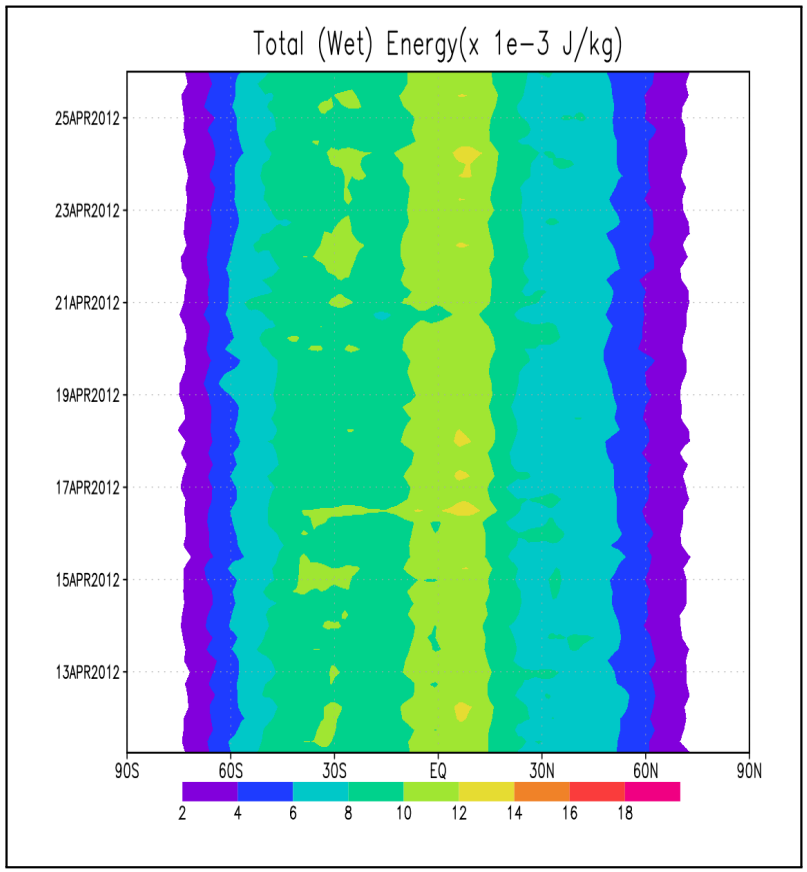


# Evolution of 6-hr Background Spread

## EnKF-based hybrid

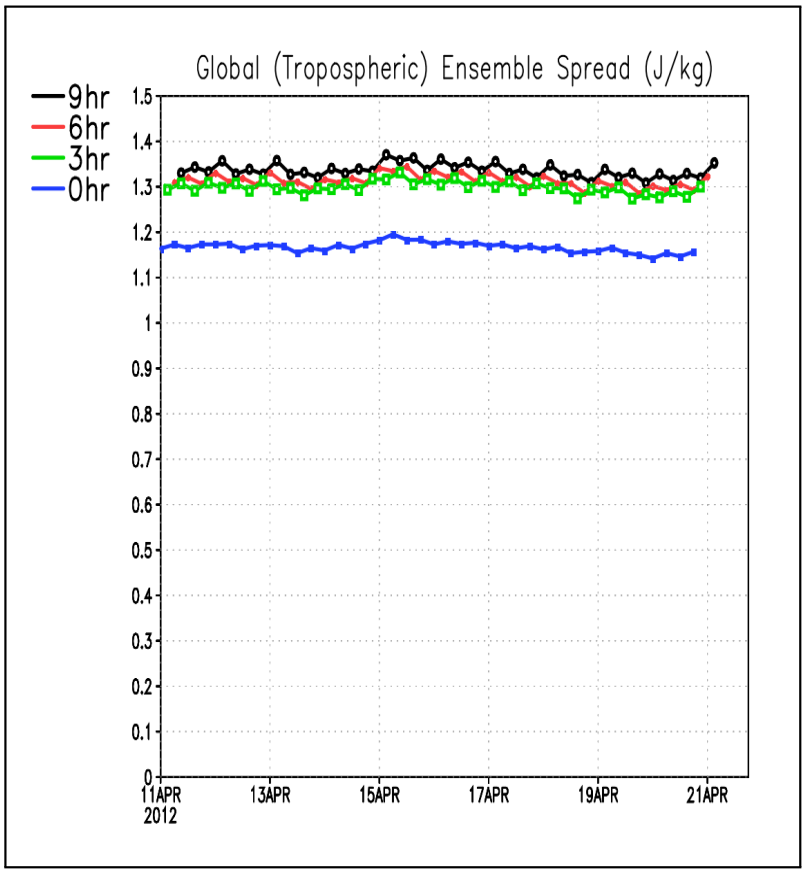



## Filter-Free hybrid

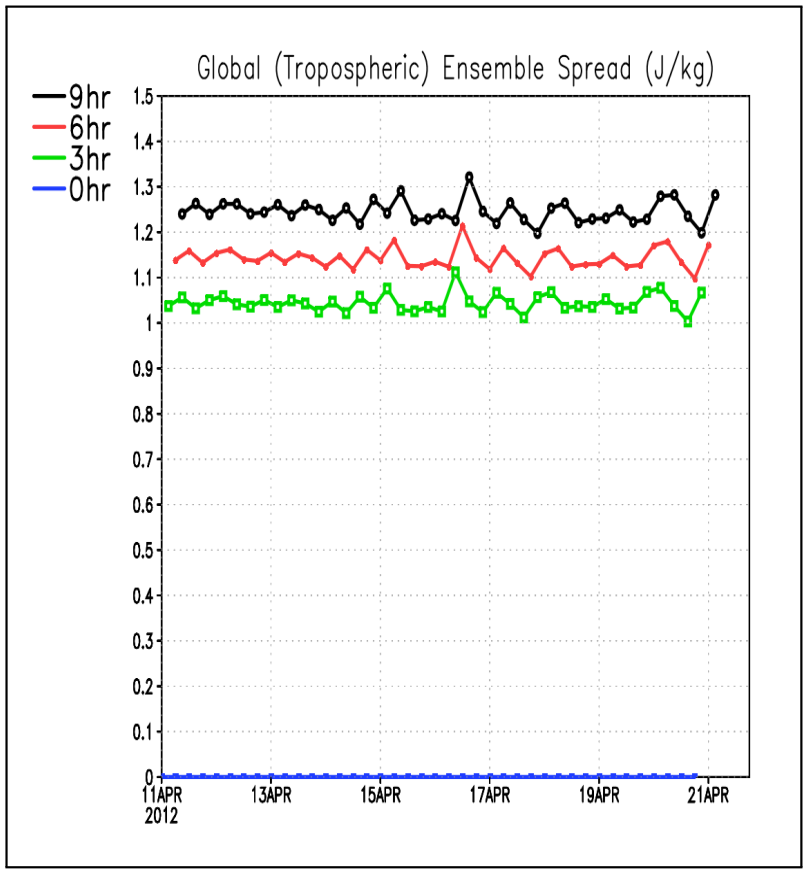


# Spread within 9-hr Background Period

## EnKF-based hybrid

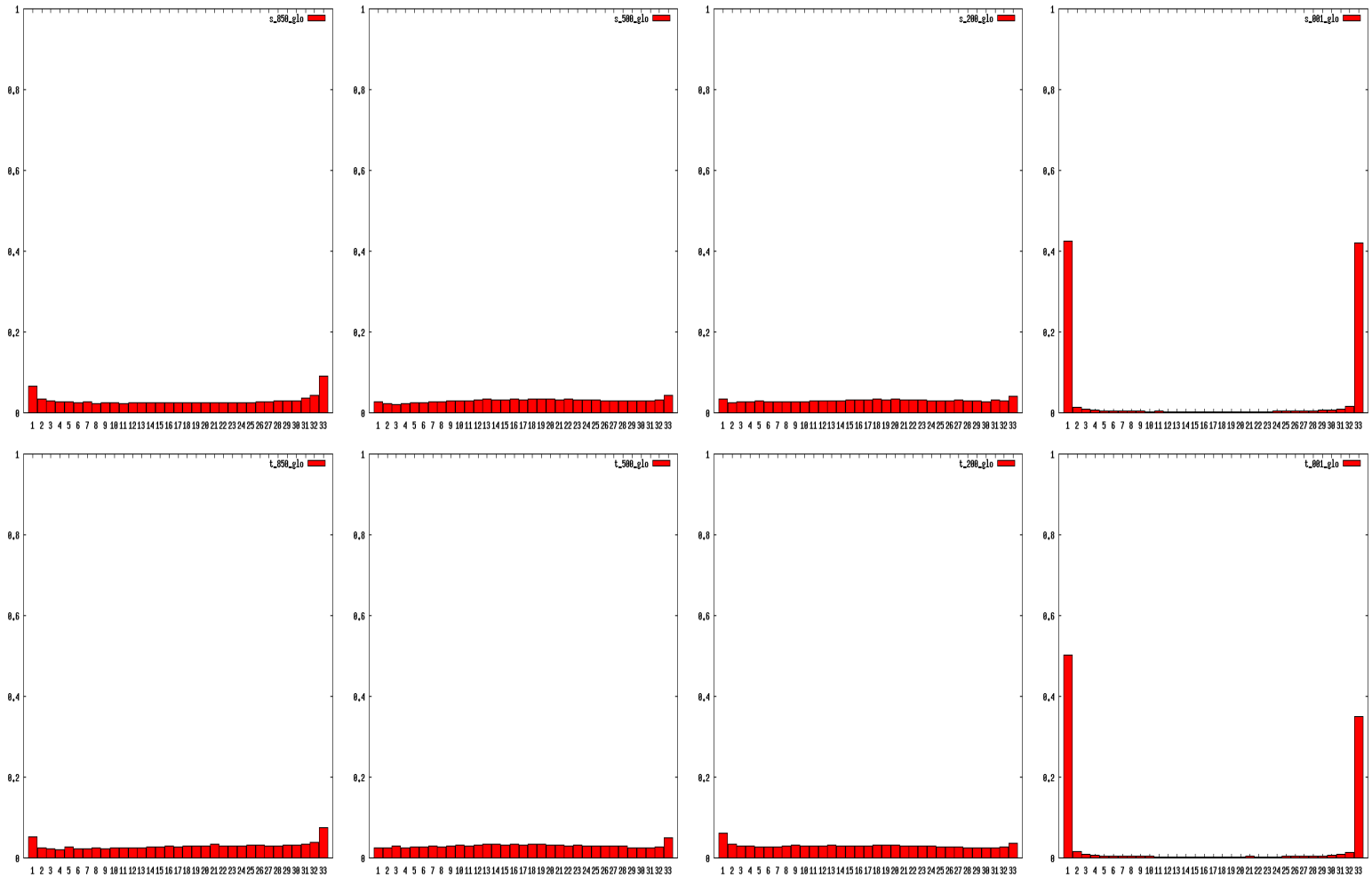



## Filter-Free hybrid



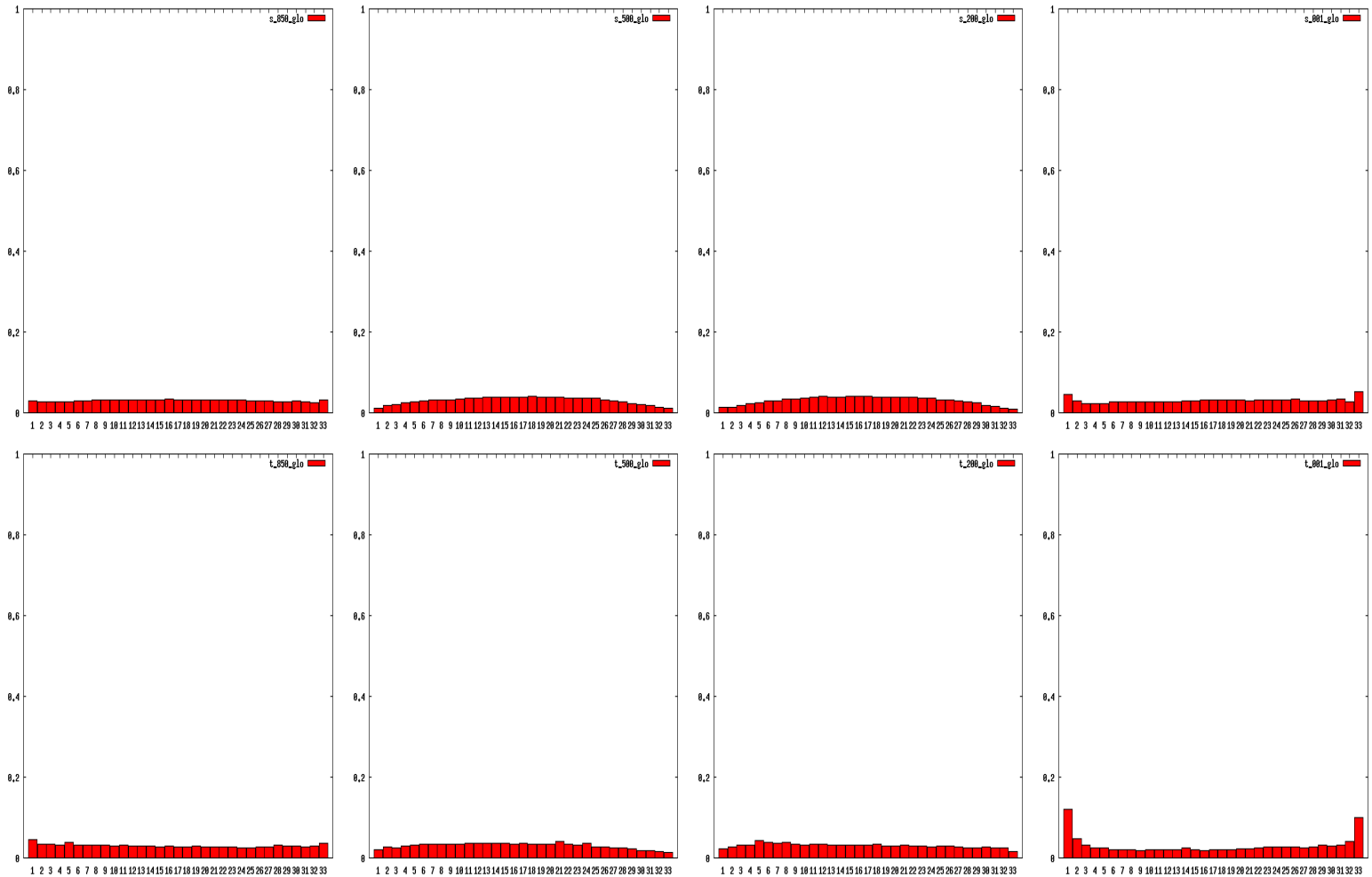
Rank-histograms, EnKF: speed (top) and temperature (bottom)

850 hPa 500 hPa 200 hPa 1 hPa



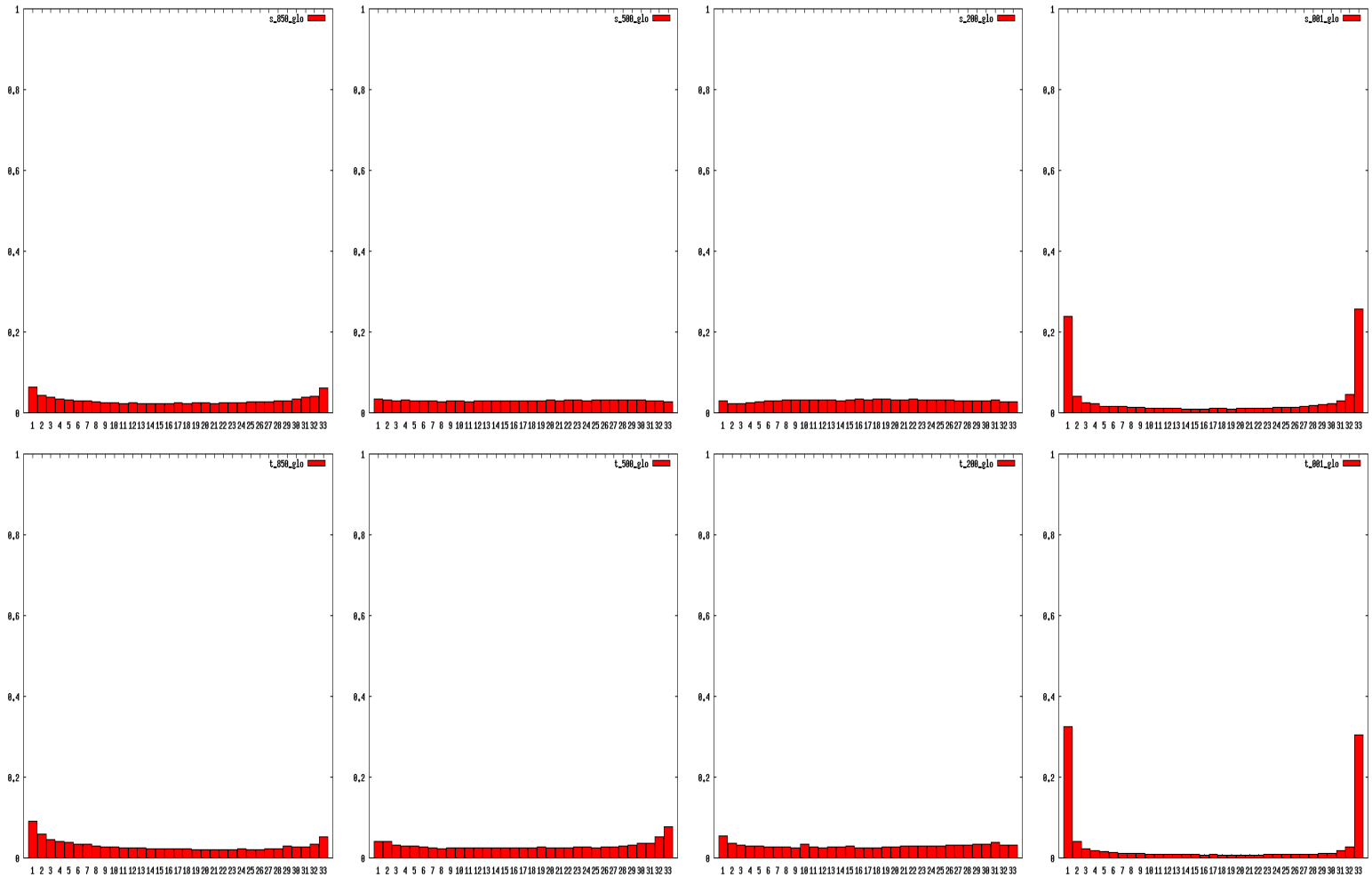
Rank-histograms, Filter-free early tuning: speed (top) and temperature (bottom)

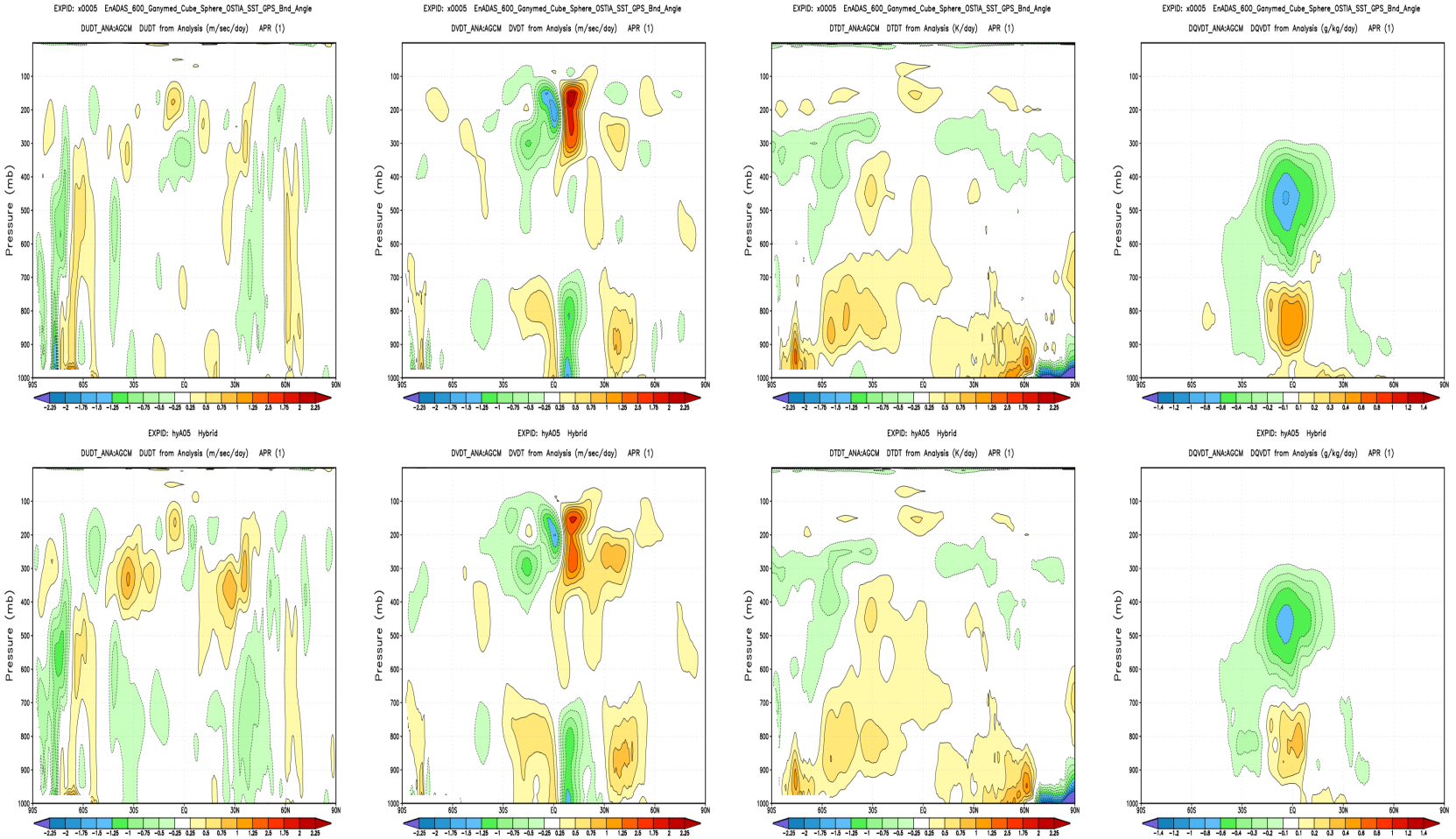
850 hPa 500 hPa 200 hPa 1 hPa



Rank-histograms, Filter-free later tuning: speed (top) and temperature (bottom)

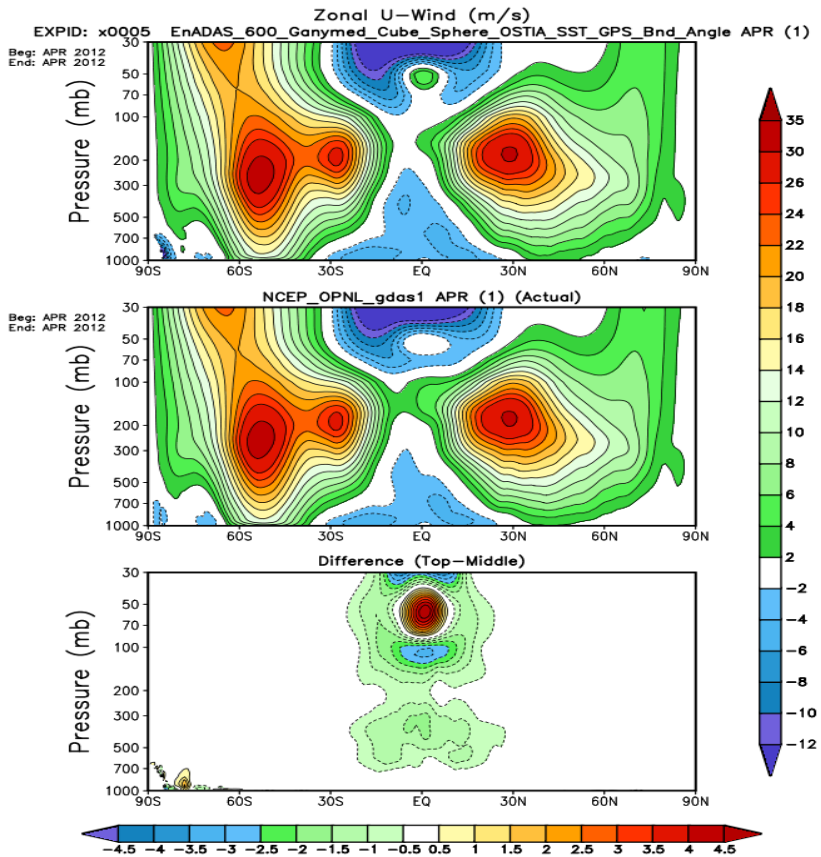
850 hPa 500 hPa 200 hPa 1 hPa



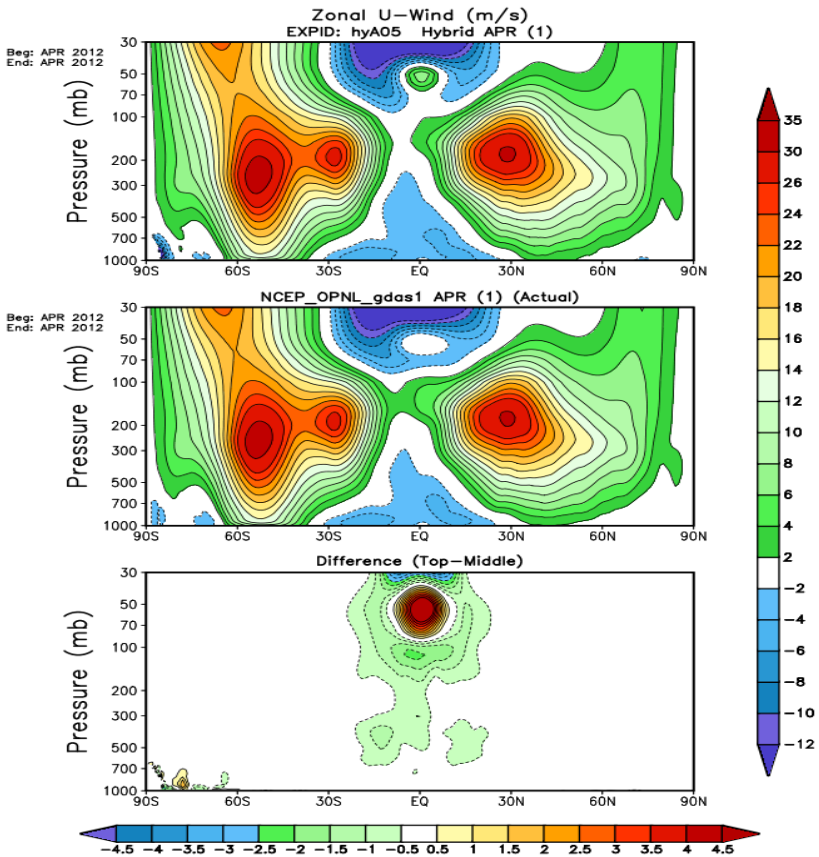

EXPID: x0005 EnADAS_600_Ganymed_Cube_Sphere_OSTIA_SST_GPS_Bnd_Angle
DUDT_ANA:AGCM DUDT from Analysis (m/sec/day) APR (1)
DVDT_ANA:AGCM DVDT from Analysis (m/sec/day) APR (1)
DTDT_ANA:AGCM DTDT from Analysis (K/day) APR (1)
DQVDT_ANA:AGCM DQVDT from Analysis (g/kg/day) APR (1)
EXPID: hyA05 Hybrid
Pressure (mb)
90S 60S 30S EQ 30N 60N 90N

# Comparison w/ NCEP: Zonally-Averaged Monthly Mean U-Wind

## Control 3d-Var

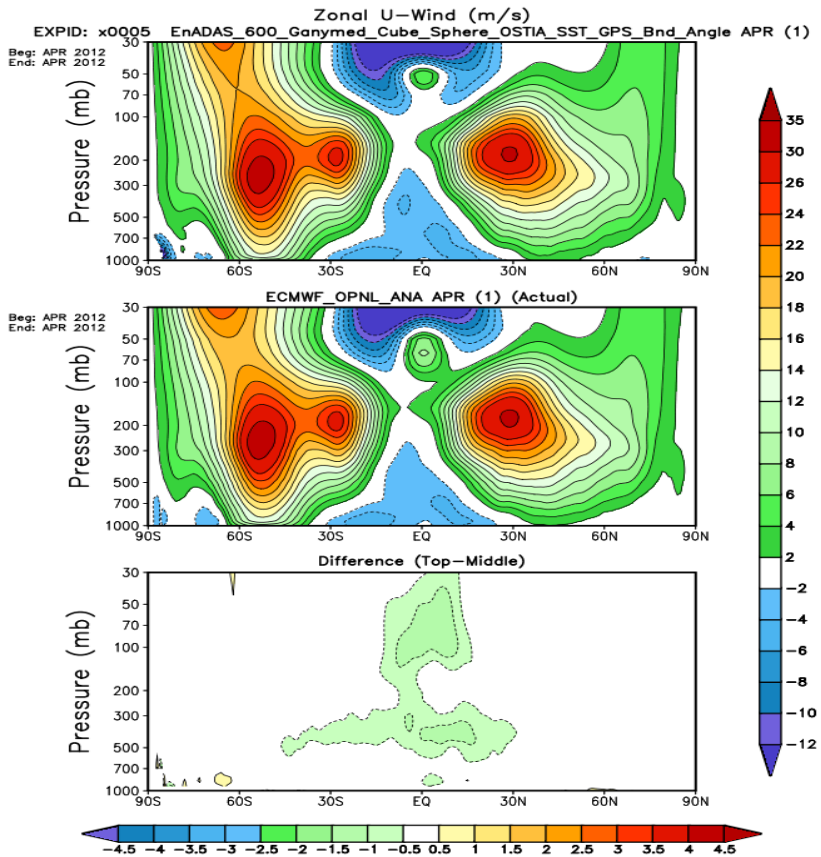



## EnKF-based hybrid

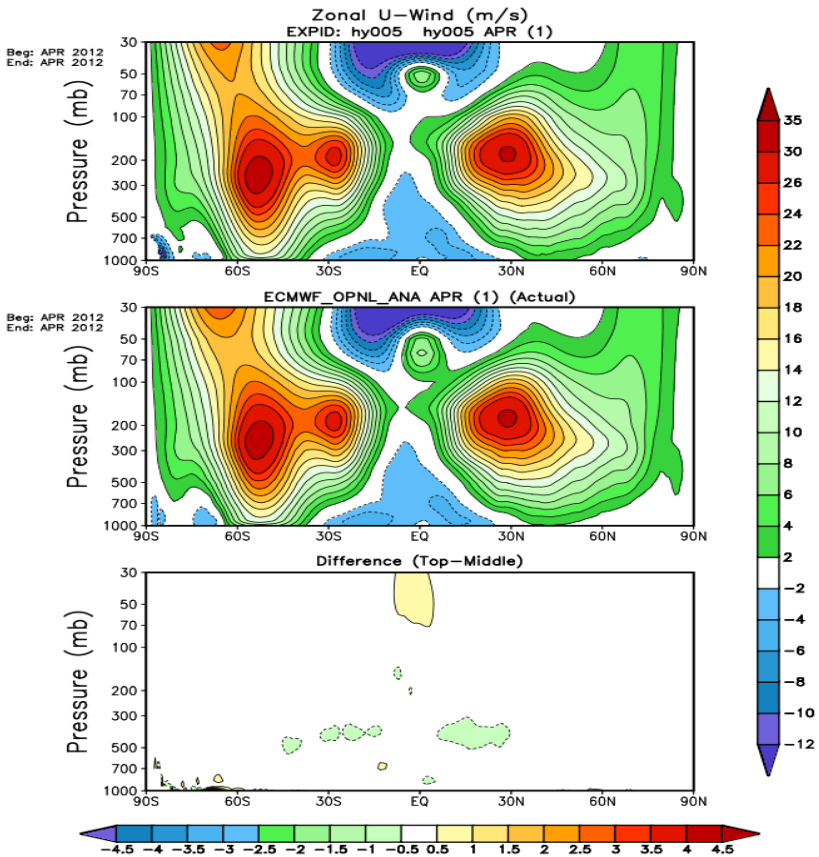


# Comparison w/ ECMWF: Zonally-Averaged Monthly Mean U-Wind

## Control 3d-Var

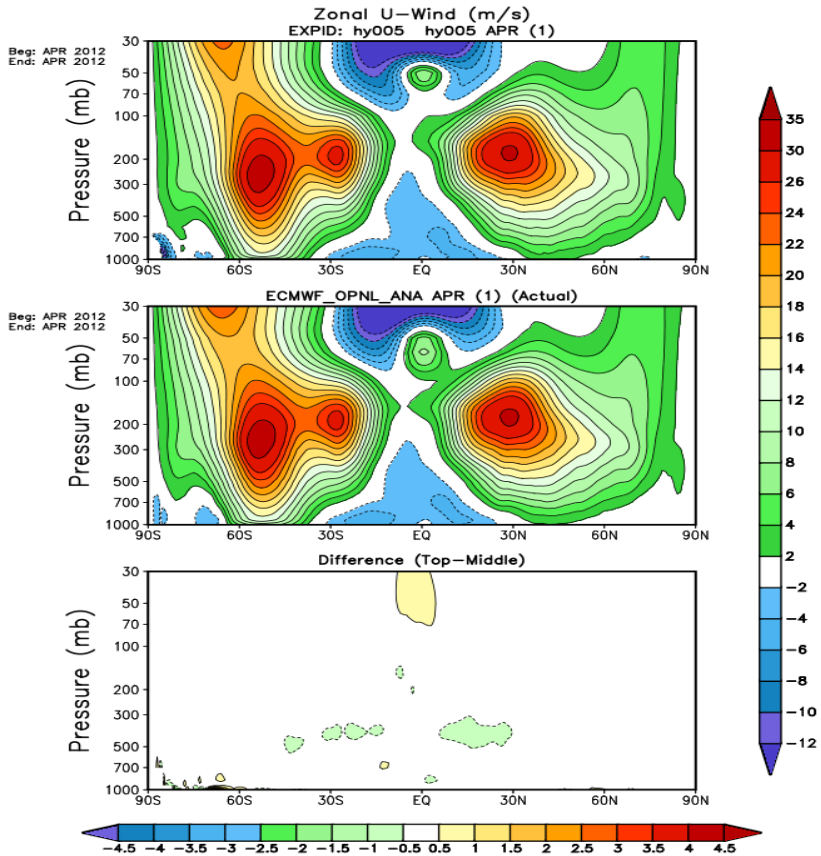



## EnKF-based hybrid

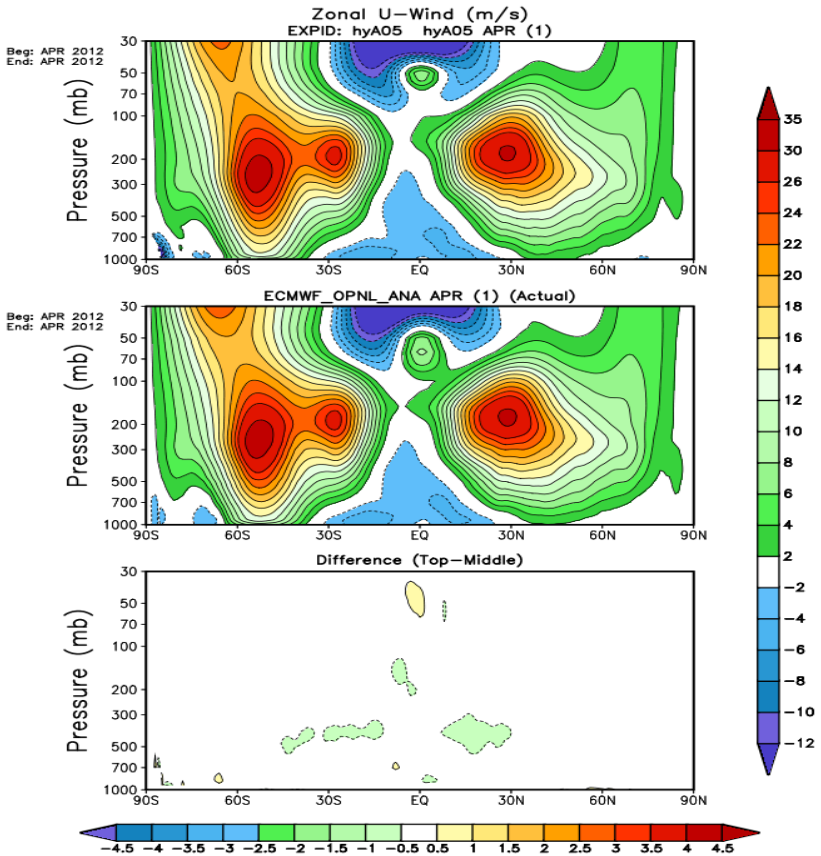


# Comparison w/ ECMWF: Zonally-Averaged Monthly Mean U-Wind

## EnKF-based hybrid

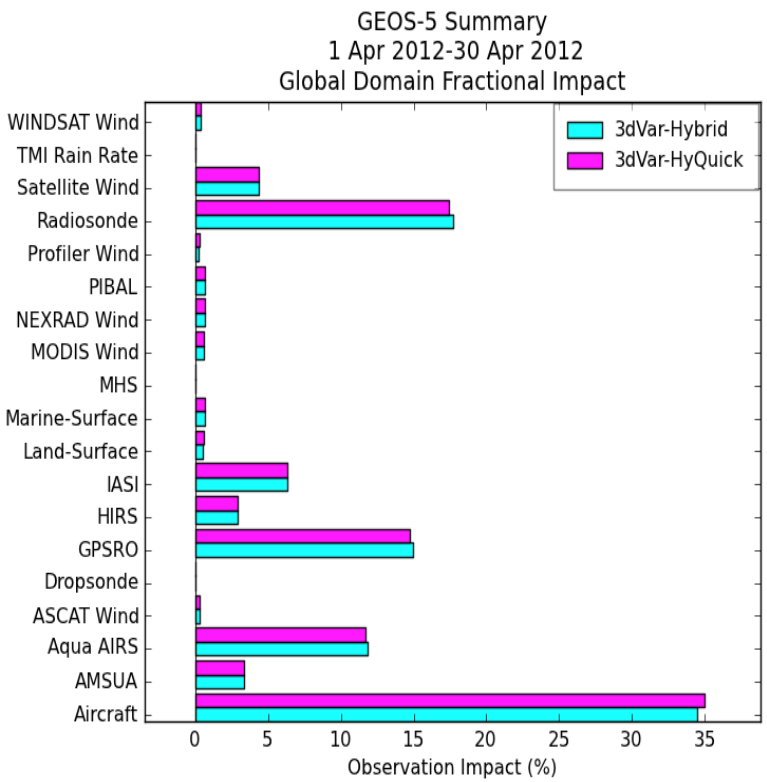



## Filter-Free hybrid

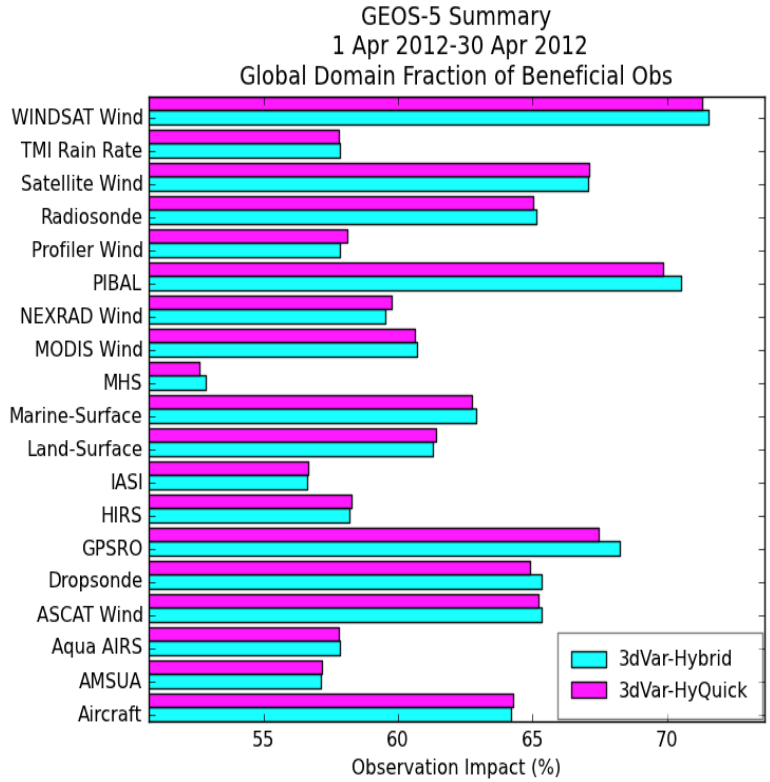


# Observation Impact on Analysis

## Fractional



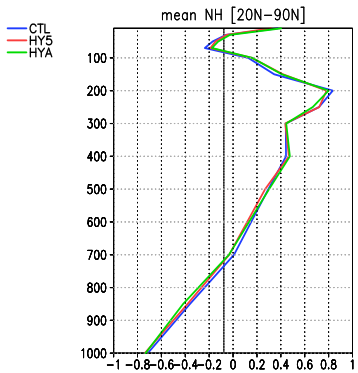

## Beneficial

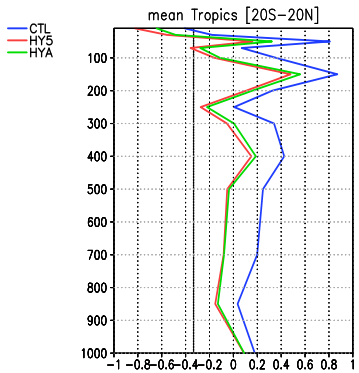


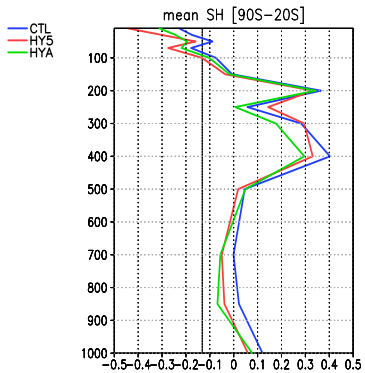
# Observations fit to background

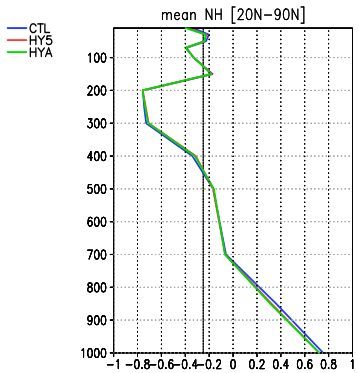
## Raob Biases: Zonal Winds (top); Temperature (bottom)

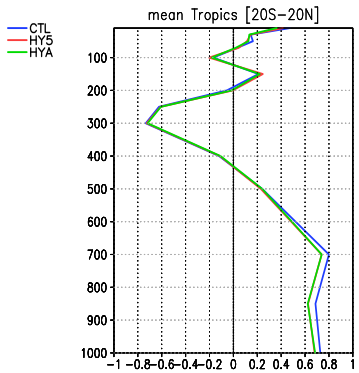
NH Tropics SH

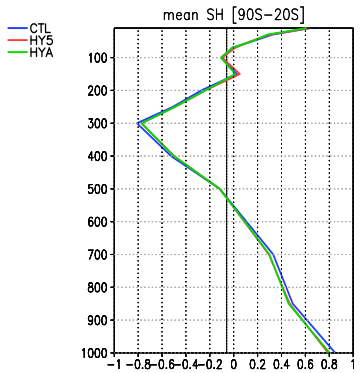


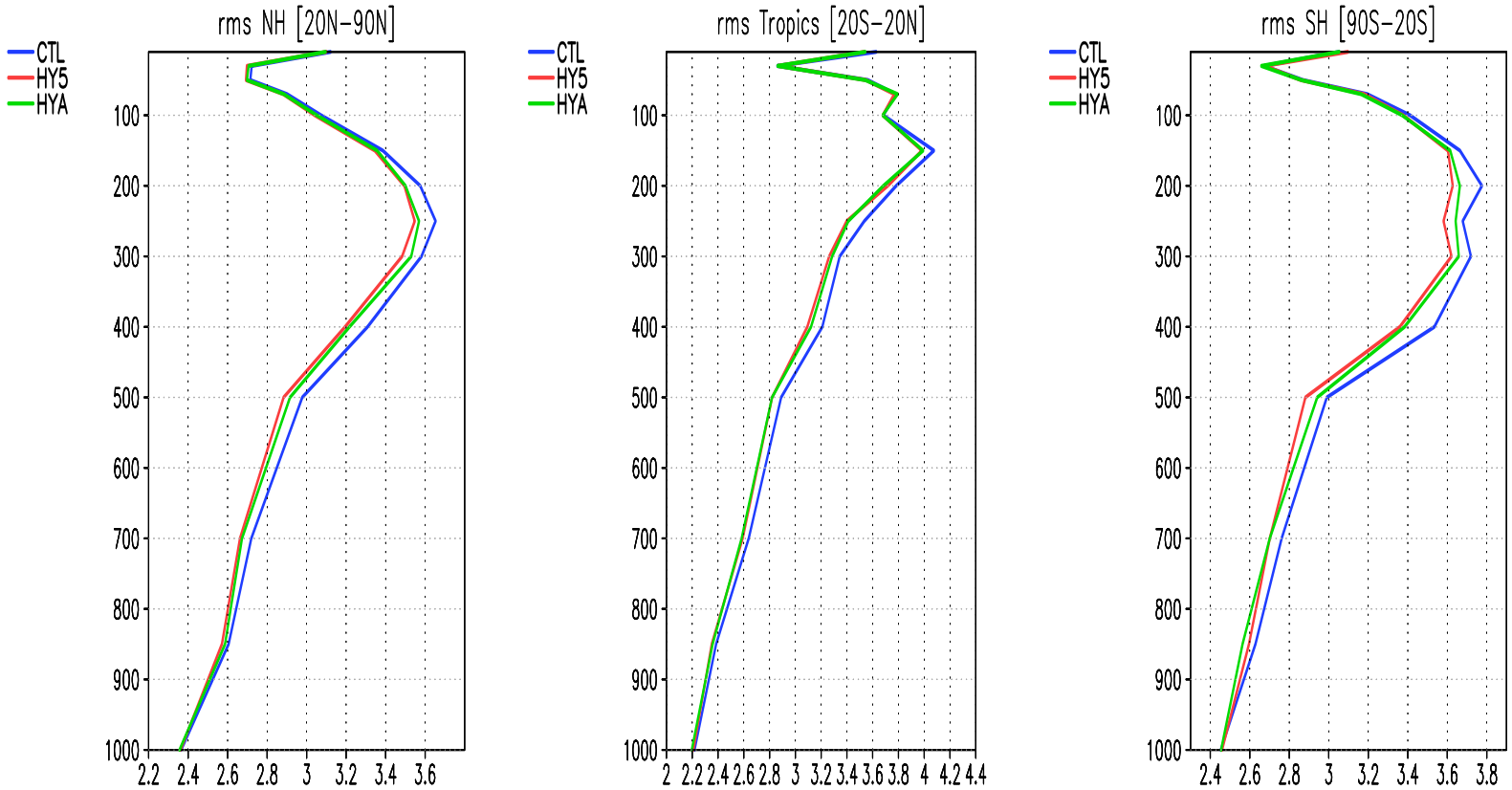



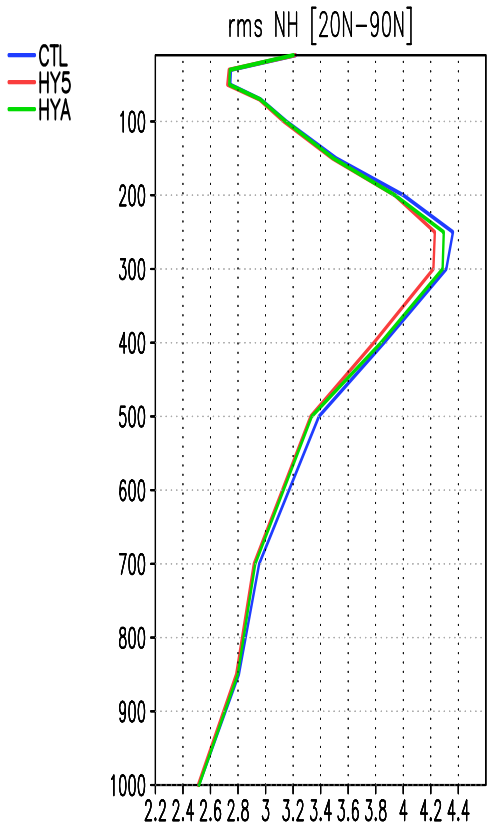
# Observations fit to background

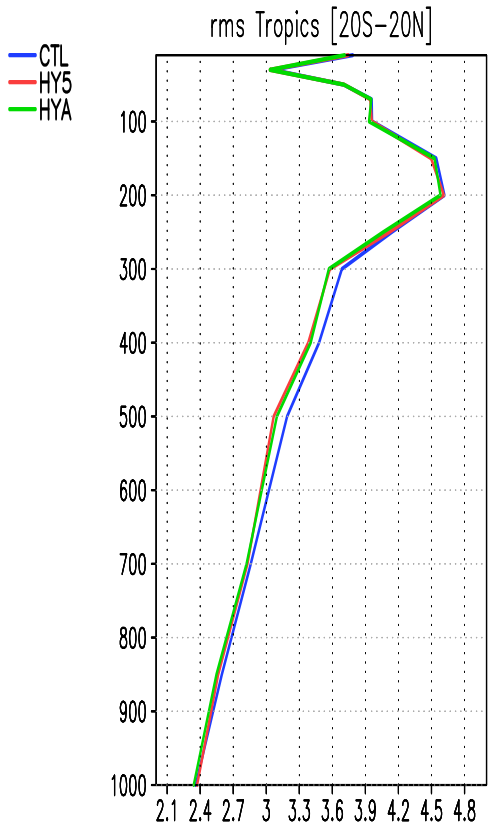
## Raob Zonal Winds RMS

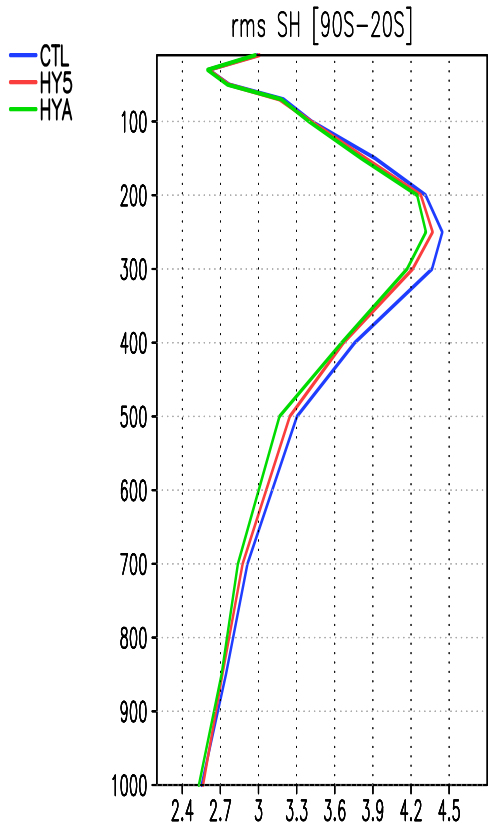


# Observations fit to 24-hour forecast

## Raob Zonal Winds

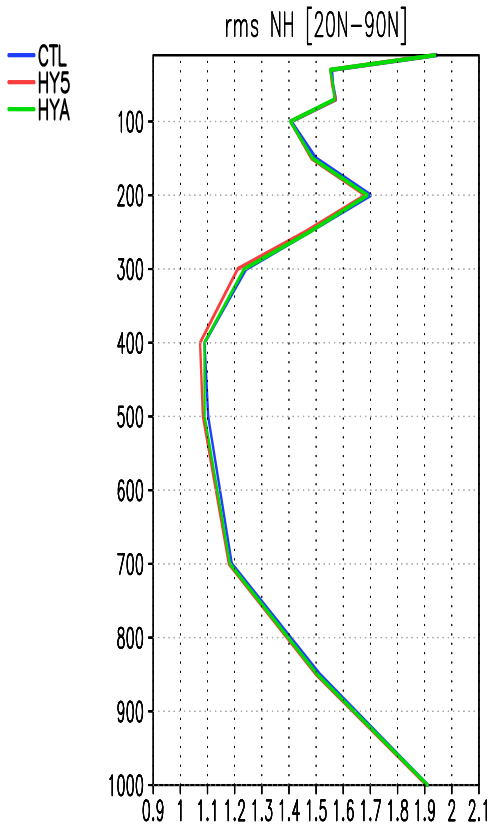


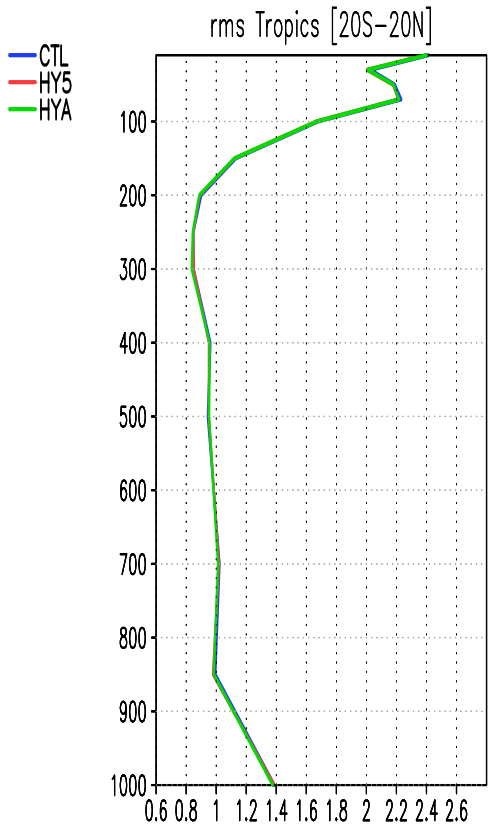



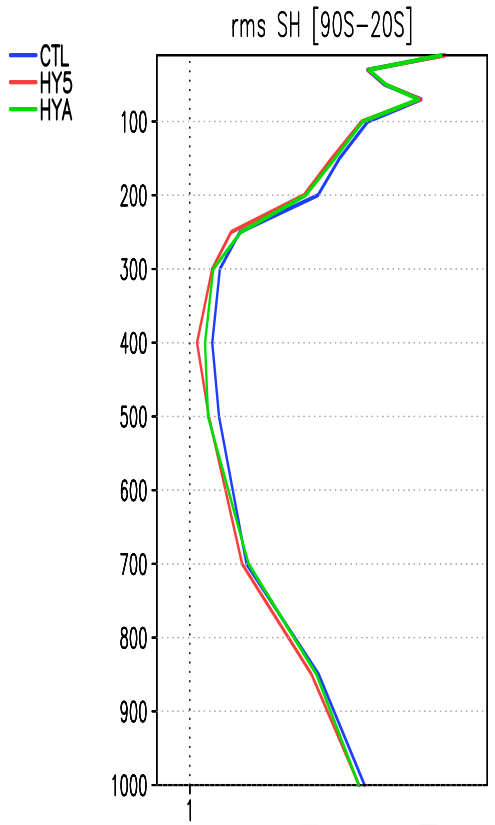



## 24-hour

## Raob Temperatures

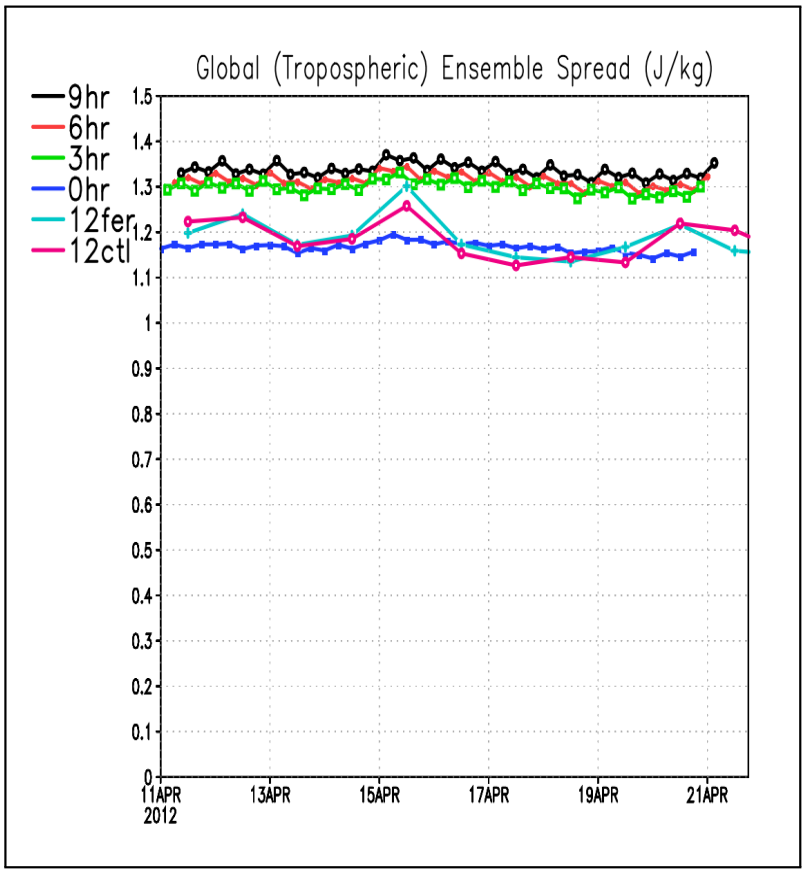


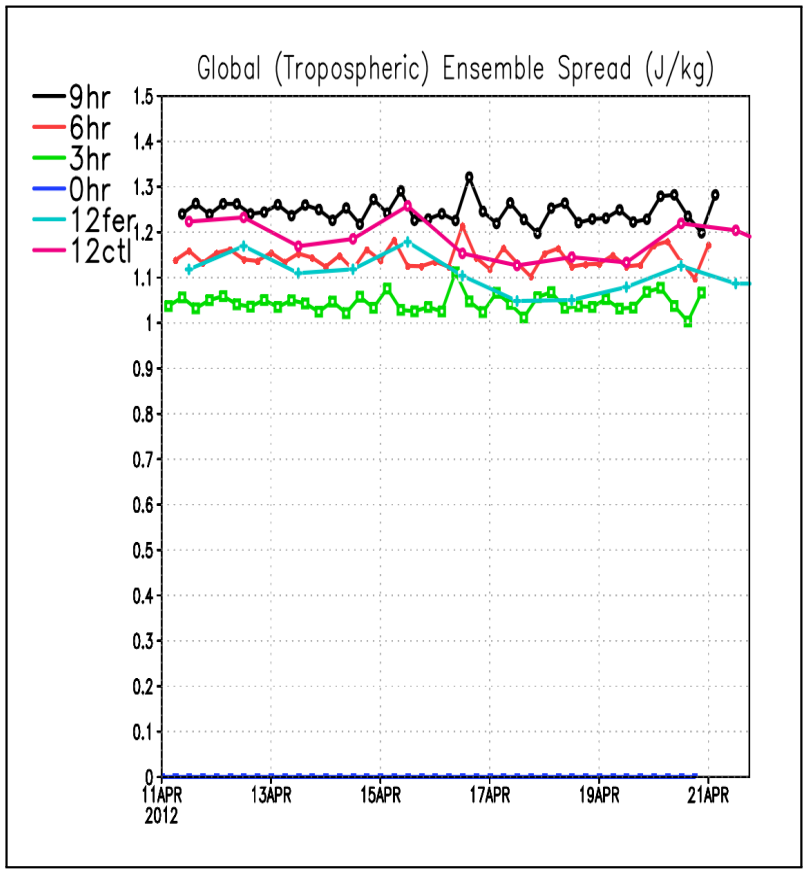



# 12-hour Error and Ensemble Spread within Background Period

## EnKF-based hybrid

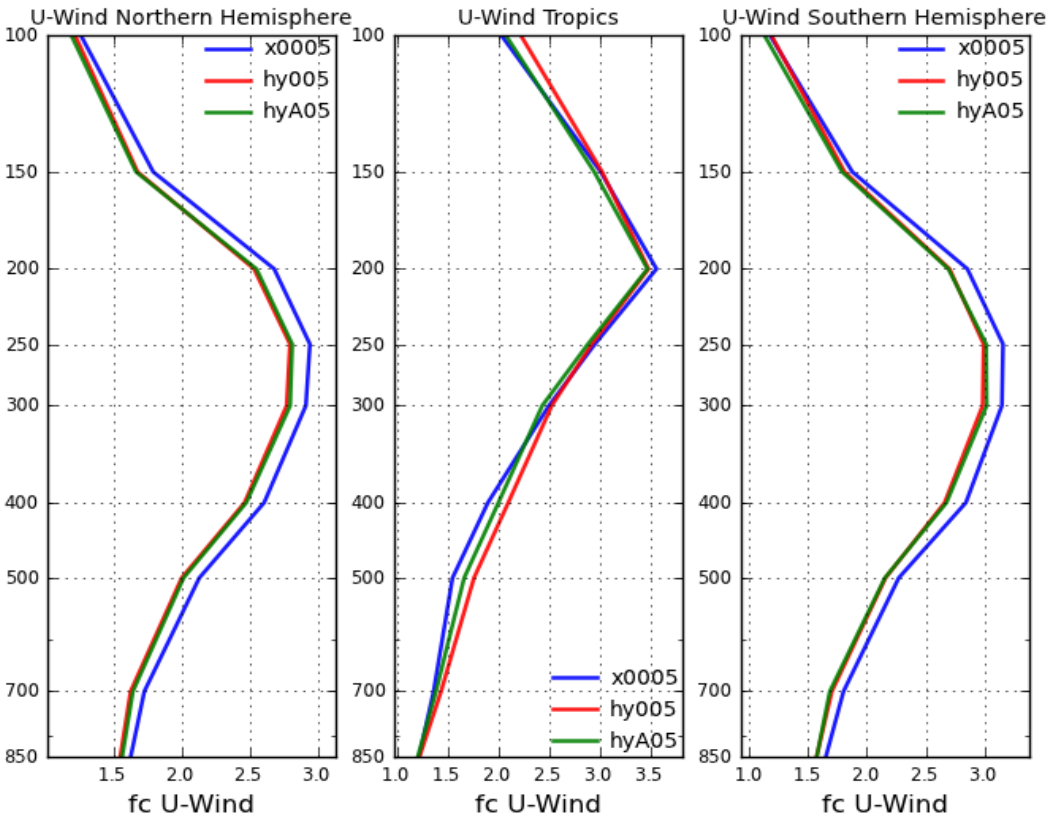



## Filter-Free hybrid



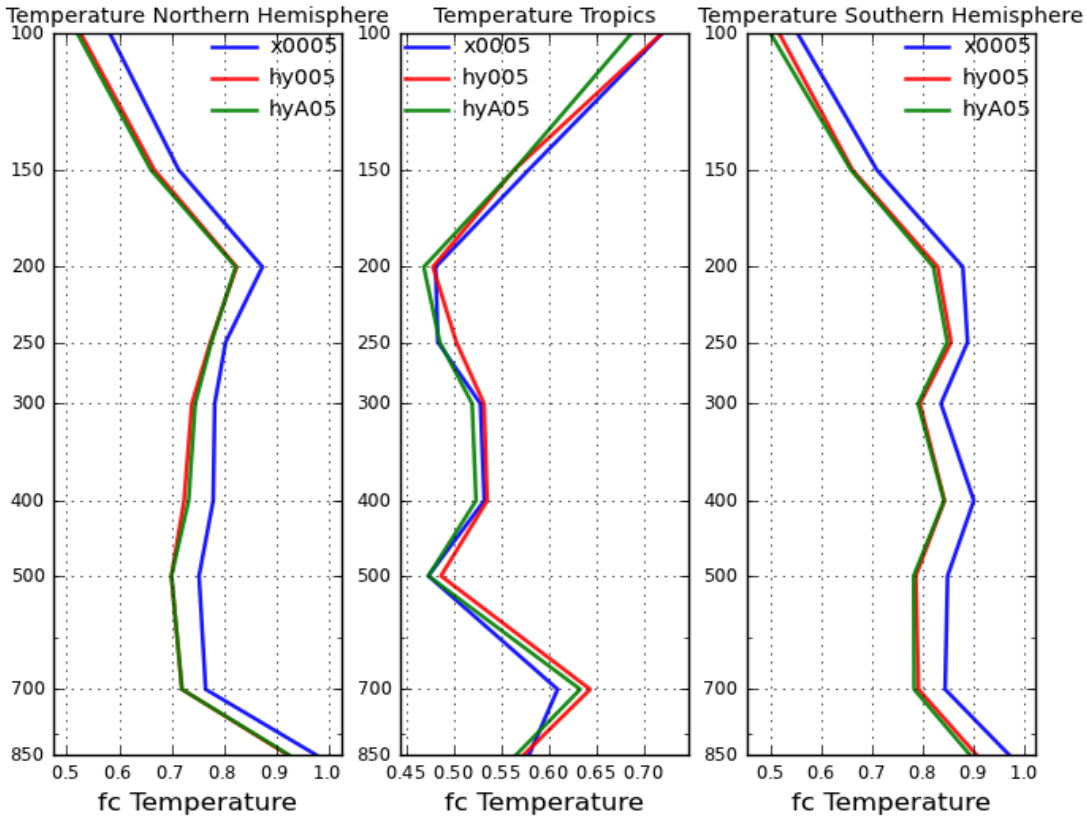
# 24-hour: Zonal Winds

NH Tropics SH



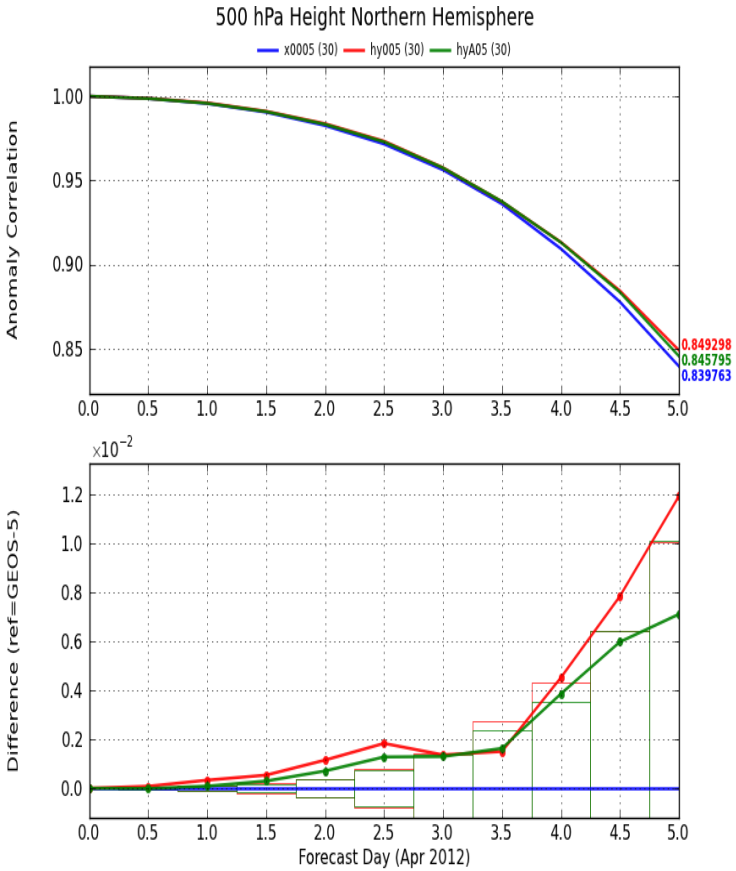
# 24-hour: Temperature

NH Tropics SH

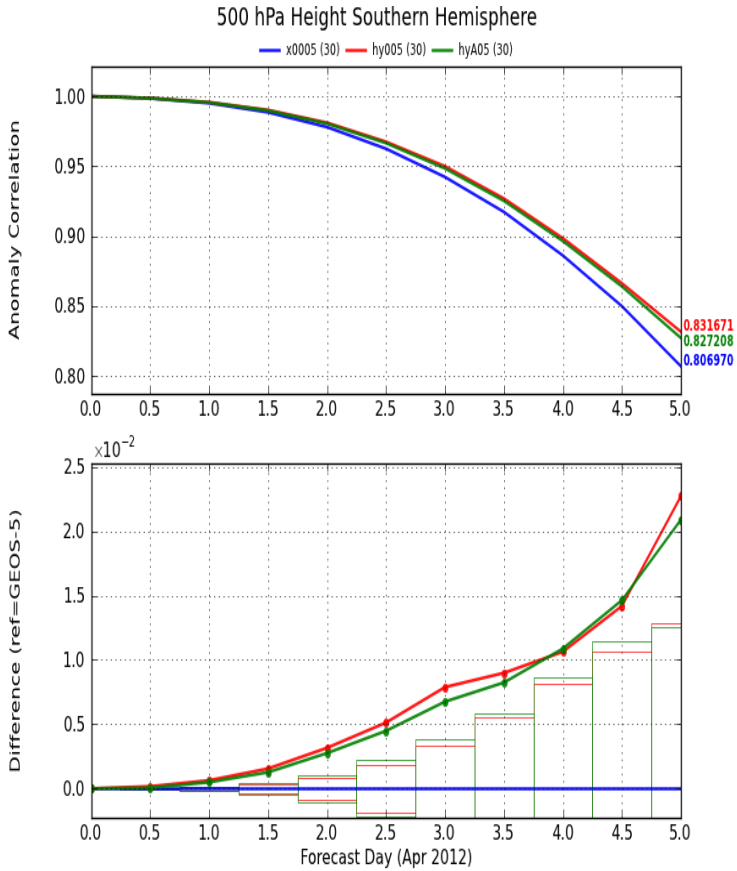


# Anomaly Correlations: H500

## Northern Hemisphere




## Southern Hemisphere



# Summary

## Main Points

- Overall 3d-hybrid approach gives positive results in GEOS DAS with noticeable reduction of model biases and improved skill scores
- Filter-free scheme works just as well as EnKF in sustaining ensemble
- Would be nice to study skill of NMC-like perturbations in an EPS

## Advantages of Filter-Free Hybrid

- Really inexpensive way of generating ensemble
- Avoids need to maintain two analysis systems
- Avoids contradictions when calculating adjoint-based obs impact

More tests taking place before deciding on operational scheme

*This work benefited tremendously from continual collaboration with NCEP*